# Supplementary material

## Collagen peptides promote skin collagen synthesis by modulating gut microbiota and activating the TGF-β pathway

Haowen Zhang[a, b], Zongliang Yao[c], Yang Song[b], Qinglian Hua[d], Xin Geng[b], Fan Zhou[b], Qingcui Li[b], Zuozhen Li[e], Zhen Luo[e], Jin Sun[a, b], Ce Qi[b*], Duo Li[a, b*]

[a] School of Pharmacy, Qingdao University, Qingdao 266071, China

[b] Institute of Nutrition and Health, Qingdao University, Qingdao 266071, China

[c] The Staff Hospital of Qingdao University, Qingdao University, Qingdao 266071, China

[d] College of Food Science and Engineering, Ocean University of China, Qingdao 266000, China

[e] Infinitus Polysacchride and Immunity Research Lab, Guangzhou 510665, China

[*] Corresponding author: Ce Qi (ceqi@qdu.edu.cn); Duo Li (duoli@qdu.edu.cn)

Address: No. 308, Ningxia Road, Laoshan District, Qingdao 266071, the People's Republic of China.

## Supplementary Methods

### 1. Molecular weight distribution determination

The molecular weight distribution of peptides was analyzed using high-performance liquid chromatography (HPLC) on a Shimadzu LC-20A system (Shimadzu, Kyoto, Japan) equipped with a TSKgel G2000 SWXL column (7.8 × 300 mm, TOSOH Bioscience, Tokyo, Japan). The mobile phase consisted of a filtered and degassed mixture of water and acetonitrile (55:45, v/v), delivered at a flow rate of 0.5 mL/min. Peptide elution was monitored at an absorbance wavelength of 220 nm. The molecular weight standards employed for the calibration curve, which was constructed using the log molecular weight versus retention time method, included cytochrome C (12,588 Da), insulin (5733 Da), glutathione (307 Da), and bacitracin (1423 Da).

### 2. Amino acid composition analysis

The peptide samples were hydrolyzed in 6.0 M HCl at 110°C for 24 hours to release individual amino acids. Following hydrolysis, the amino acids were derivatized with phenylisothiocyanate (PITC) to enhance detection and facilitate analysis. The amino acid compositions were determined using reversed-phase high-performance liquid chromatography (RP-HPLC).

The mobile phase consisted of a 10 mM phosphate buffer solution with a pH of 6.9 (A) and acetonitrile (B). A gradient elution program was employed as follows: 0-5 minutes, 5-10% B; 5-25 minutes, 10-17% B; 25-45 minutes, 17-35% B; 45-48 minutes, 35-100% B; 48-50 minutes, 100% B; 50-58 minutes, 100-5% B; and 58-60 minutes, 5% B. The flow rate of the mobile phase was 1.0 mL/min and the detection wavelength was set at 254 nm.

### 3. Identification of tripeptides

Peptide samples were first mixed with 0.25% acetic acid, followed by

centrifugation to remove the supernatant. The resulting samples were transferred to 10 kDa ultrafiltration centrifugal tubes and centrifuged at 12,000 × g for 15 minutes. To further wash the peptides, 200 μL of 0.25% acetic acid was added, and the centrifugation step was repeated. This process was performed twice. The collected filtrate was desalted using a C18 StageTip, then dried under vacuum. After drying, the peptides were re-solubilized in 0.1% formic acid, and the peptide concentration was determined by measuring the absorbance at 280 nm (OD280) before subjecting the sample to LC-MS/MS analysis.

For chromatographic separation, an appropriate amount of peptides from each sample was loaded onto a nanoliter flow rate Easy nLC 1200 chromatography system (Thermo Scientific). The buffers used for chromatographic separation were as follows: solution A consisting of 0.1% formic acid in water, and solution B consisting of 0.1% formic acid, acetonitrile, and water (with acetonitrile at 80%). The chromatographic column was first equilibrated with 100% solution A. The samples were injected onto a Trap Column (100 μm × 20 mm, 5 μm, C18, Dr. Maisch GmbH) and subsequently passed through a C18 analytical column (75 μm × 150 mm, 3 μm, Dr. Maisch GmbH) for gradient separation at a flow rate of 300 nL/min. The gradient for liquid-phase separation was as follows: 0–2 min, liquid B increased linearly from 2% to 5%; 2–44 min, liquid B increased linearly from 5% to 28%; 44–51 min, liquid B increased linearly from 28% to 40%; 51–53 min, liquid B increased linearly from 40% to 100%; and 53–60 min, liquid B was maintained at 100%.

The peptides were then separated and analyzed by data-dependent acquisition (DDA) mass spectrometry using a Q-Exactive HF-X mass spectrometer (Thermo Scientific). The analysis duration was set to 60 minutes with the detection mode set to positive ions. The parent ion scan range was 50–750 m/z, with a primary mass resolution of 60,000 at m/z 200. The automatic gain control (AGC) target was set to 3e6, and the primary maximum ion trap time (IT) was set to 50 ms. Peptide fragmentation spectra (MS2) were acquired for the 20 highest intensity parent ions after each full scan. MS2 scan parameters were as follows: MS2 resolution of 15,000 at m/z 200, AGC target of 1e5, MS2 maximum IT of 50 ms, MS2 activation type: high-energy collisional dissociation (HCD), isolation window of 1.6 m/z, and normalized collision energy set to 28.

**4. Quantification of tripeptides**

Standard tripeptide solutions were prepared in ultrapure water at concentrations ranging from 1.6 to 1000 μg/mL. Peptide samples were dissolved in 10% aqueous acetonitrile to achieve a final concentration of 5.0 mg/mL. The solutions were then filtered through a 0.22-μm microporous membrane and stored for subsequent analysis. Plasma and gut sac fluid samples were processed by vortex-mixing 50 μL of the sample with 250 μL of pre-chilled acetonitrile, followed by incubation on ice for 15 minutes. The mixture was centrifuged at 15,000 rpm for 15 minutes, and the resulting supernatant was filtered through a membrane for further analysis.

Tripeptide contents were quantified using an ACQUITY UPLC system (Waters, Milford, MA, USA) coupled with electrospray ionization (ESI) and an Impact II quadrupole time-of-flight (Q-TOF) mass spectrometer (Bruker Daltonics, Bremen,

Germany). Briefly, chromatographic separation was performed on a Waters ACQUITY UPLC HSS T3 column (2.1 mm × 100 mm, 1.8 μm, 100 Å). The injection volume of the prepared sample was 2 μL. The flow rate was maintained at 200 μL/min, and the column temperature was set at 30°C. The mobile phases consisted of 0.1% formic acid in ultrapure water (A) and acetonitrile (B). Gradient elution conditions were as follows: 0–10 min, 0%–30% B; 10–11 min, 30%–85% B; 11–13 min, 85% B; 13–13.5 min, 85%–0% B; 13.5–17 min, 0% B.

**5. RNA sequencing and gene set enrichment analysis for skin**

Total RNA was extracted from skin samples using the RNeasy Kit (Qiagen, Hilden, Germany). RNA concentration and purity were assessed using a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Wilmington, USA), while RNA integrity was evaluated with the RNA Nano 6000 Assay Kit on the Agilent Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, USA). High-quality RNA was subsequently used for cDNA library construction and sequencing, conducted at Biomarker Technologies Corporation (Beijing, China). cDNA libraries were prepared using the NEBNext Ultra™ RNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, USA) following the manufacturer's protocol. Briefly, mRNA was isolated using the NEBNext Poly(A) mRNA Magnetic Isolation Module and subsequently fragmented. First-strand cDNA synthesis was performed, followed by the generation of double-stranded cDNA using DNA Polymerase I and RNase H. The cDNA products were purified using the Agencourt AMPure XP system (Beckman Coulter, Brea, USA), subjected to end repair, and ligated with sequencing adapters. PCR amplification enriched the adapter-ligated products, which were again purified using the AMPure XP system. Sequencing was performed on the Illumina HiSeq 2500 platform. Raw sequencing reads were processed using custom Perl scripts to remove adapter sequences, filter out reads containing more than 5% unknown nucleotides, and eliminate low-quality reads. High-quality clean reads were aligned to the mouse genome (mm10) using TopHat2. Gene expression levels were quantified using fragments per kilobase of transcript per million mapped reads (FPKM).

Differential gene expression analysis was conducted using the R package *limma*, employing Bayesian-adjusted t-statistics from linear models. Multiple testing correction was performed using the false discovery rate (FDR). DEGs were identified based on log2 fold-change (log2FC) greater than 1.5 and an FDR less than 0.05. DEGs meeting the specified criteria were subjected to Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis using the R package *clusterProfiler*. A corrected *p*-value of less than 0.05 was set as the threshold for significance. KEGG pathway visualizations were generated using the *Pathview* R package. The relative abundance of immune cells in the skin was estimated using the CIBERSORT algorithm. Protein-protein interaction (PPI) networks were constructed by mapping DEGs to the STRING database (version 12.0).

**6. 16S rRNA amplicon sequencing of cecal contents**

Bacterial DNA was extracted from cecal fecal samples using the QIAamp Fast DNA Stool Mini Kit (Qiagen, Cat# 51604) according to the manufacturer's instructions. PCR amplification targeted the variable V3–V4 region of the bacterial 16S

rRNA gene using barcoded primers: forward primer 5'-ACTCCTACGGGAGGCAGCA-3' and reverse primer 5'-GGACTACHVGGGTWTCTAAT-3'. Sequencing libraries were constructed, and paired-end sequencing was performed on the Illumina MiSeq platform by Biomarker Technologies Co., Ltd. (Beijing, China). Downstream bioinformatics analysis of the amplicons was conducted using QIIME 2. The DADA2 plugin was utilized to denoise sequences into amplicon sequence variants (ASVs). These ASVs were clustered into operational taxonomic units (OTUs) at 97% similarity using the vsearch cluster-features-de-novo plugin in QIIME 2.

**7. Determination of short-chain fatty acids (SCFAs) in cecal contents**

Approximately 30 mg of cecal contents were weighed into a homogenization tube, and 270 μL of deionized water was added. The mixture was homogenized using a TissueRuptor (60 Hz for 60 s), followed by centrifugation at 10,000 rpm for 10 minutes. The resulting supernatant was then filtered through a 0.45 μm nylon membrane. A 100 μL aliquot of the supernatant was taken, to which 100 μL of acetone, 10 μL of 50% sulfuric acid, 100 μL of ether, and 30 mg of sodium chloride were added. The mixture was thoroughly vortexed for 2 minutes and subsequently centrifuged at 3,000 rpm for 3 minutes. The supernatant was collected and analyzed by gas chromatography-mass spectrometry (GC-MS) (Agilent 7890B-5977B GC/MSD, USA) for SCFA detection. For quantification, a standard curve was established by preparing a gradient mix of SCFA standard solutions, which were diluted with ultrapure water and used to generate the standard curve.

**8. Untargeted metabolomics of cecal contents**

For the metabolomics analysis, 20 mg of cecal contents were accurately weighed and homogenized with 200 μL of 80% methanol and three medium-sized steel balls in a tissue homogenizer. Subsequently, 800 μL of 80% methanol was added, and the mixture was vortexed to ensure thorough mixing. The sample was then subjected to ultrasonic treatment for 20 minutes in an ice bath, followed by incubation at -20°C for 2 hours. After the incubation period, the sample was centrifuged at 16,000 × g for 20 minutes at 4°C. The resulting supernatant was evaporated to dryness using a high-speed vacuum concentrator. 40 μL of 50% methanol was added to the dried supernatant, and the mixture was centrifuged at 20,000 × g for 15 minutes at 4°C. The supernatant was then carefully collected for analysis. A 10 μL aliquot of each sample was pooled together for quality control purposes.

For chromatographic separation, all samples were stored in an autosampler at 4°C to maintain stability. Chromatographic separation was performed using an ACQUITY UPLC® HSS T3 column (2.1 × 100 mm, 1.8 μm; Waters, Milford, MA, USA) on a Shimadzu LC-30 ultra-high performance liquid chromatography (UHPLC) system. The injection volume was 4 μL, and the column temperature was maintained at 40°C. The mobile phase flow rate was set to 0.3 mL/min. The mobile phases consisted of A: 0.1% aqueous formic acid and B: acetonitrile. The gradient elution program followed the schedule below:0–2 min: 0% B; 2–6 min: B increased linearly from 0% to 48%; 6–10 min: B increased linearly from 48% to 100%; 10–12 min: 100% B; 12–12.1 min: 100% B; 12.1–15 min: 0% B. The total experiment duration was 15 minutes.

For mass spectrometry, samples were analyzed using electrospray ionization (ESI) in both positive (+) and negative (-) modes. Separation was achieved via UPLC, and detection was performed with a QE Plus mass spectrometer (Thermo Scientific) equipped with a heated electrospray ionization (HESI) source. The ionization conditions were as follows: spray voltage of 3.8 kV (+) and 3.2 kV (-), capillary temperature set at 320°C (±), sheath gas flow at 30 (±), auxiliary gas flow at 5 (±), probe heater temperature at 350°C (±), and the S-lens RF level set to 50. The MS acquisition time was 15 minutes, with a parent ion scanning range of 70–1050 m/z. The primary mass resolution was 70,000 at m/z 200, and the automatic gain control (AGC) target was set to 3e6, with a maximum ion trap time (IT) of 100 ms. Secondary mass spectrometry (MS2) was performed by selecting the 10 most intense parent ions for fragmentation. The MS2 scan parameters were as follows: resolution of 15,000 at m/z 200, AGC target of 1e5, and a maximum IT of 50 ms. The MS2 activation type was high-energy collisional dissociation (HCD), with an isolation window of 2 m/z. The normalized collision energy was set in steps at 20, 30, and 40.

## Supplementary Tables

**Table S1.** Molecular weight distributions of peptides.

| | Relative content (%) | | | |
|---|---|---|---|---|
| Sample | 200–500 Da | 500–1000 Da | 1000–2000 Da | >2000 Da |
| CTP | 49.14 ± 0.36 | 27.72 ± 0.17 | 15.31 ± 0.36 | 7.83 ± 0.27 |
| HCP | 11.08 ± 0.27 | 20.99 ± 0.11 | 22.41 ± 0.15 | 45.53 ± 0.09 |
| CAP | 7.03 ± 0.06 | 28.40 ± 0.17 | 32.04 ± 0.06 | 32.53 ± 0.15 |

CTP, hydroxyproline-containing tripeptide-rich collagen peptides; HCP, high molecular weight collagen peptides; CAP, casein peptides. Data are shown as the mean ± SEM (n = 3 per group).

**Table S2.** Amino acid compositions of peptides.

| | Content (%) | | |
|---|---|---|---|
| Amino acid | CTP | HCP | CAP |
| Asp | 5.51 ± 0.03 | 5.26 ± 0.04 | 6.94 ± 0.05 |
| His | 0.61 ± 0.01 | 0.83 ± 0.01 | 2.56 ± 0.01 |
| Arg | 7.96 ± 0.02 | 8.50 ± 0.01 | 3.02 ± 0.03 |
| Lys | 3.52 ± 0.01 | 3.42 ± 0.02 | 7.31 ± 0.01 |
| Ser | 3.62 ± 0.01 | 3.59 ± 0.03 | 5.52 ± 0.01 |
| Glu | 10.02 ± 0.01 | 9.47 ± 0.01 | 20.86 ± 0.08 |
| Gly | 23.64 ± 0.08 | 23.49 ± 0.10 | 2.21 ± 0.02 |
| Met | 1.00 ± 0.01 | 0.78 ± 0.00 | 2.80 ± 0.01 |
| Leu | 2.37 ± 0.01 | 2.62 ± 0.02 | 9.35 ± 0.01 |
| Phe | 1.53 ± 0.01 | 1.92 ± 0.01 | 4.91 ± 0.02 |
| Thr | 2.70 ± 0.02 | 2.64 ± 0.01 | 4.20 ± 0.01 |
| Pro | 12.49 ± 0.07 | 12.76 ± 0.08 | 10.00 ± 0.08 |
| Ala | 10.34 ± 0.04 | 10.19 ± 0.06 | 4.71 ± 0.01 |
| Ile | 1.08 ± 0.01 | 1.06 ± 0.01 | 4.50 ± 0.04 |
| Val | 2.00 ± 0.01 | 1.89 ± 0.02 | 6.22 ± 0.01 |
| Tyr | 0.17 ± 0.01 | 0.34 ± 0.03 | 4.88 ± 0.03 |
| Hyp | 11.42 ± 0.10 | 11.24 ± 0.12 | 0.01 ± 0.00 |

Amino acids are represented using the three-letter amino acid code. Hyp, hydroxyproline; CTP, Hyp-containing tripeptide-rich collagen peptides; HCP, high molecular weight collagen peptides; CAP, casein peptides. Data are shown as the mean ± SEM (n = 3 per group).

## Supplementary Figures

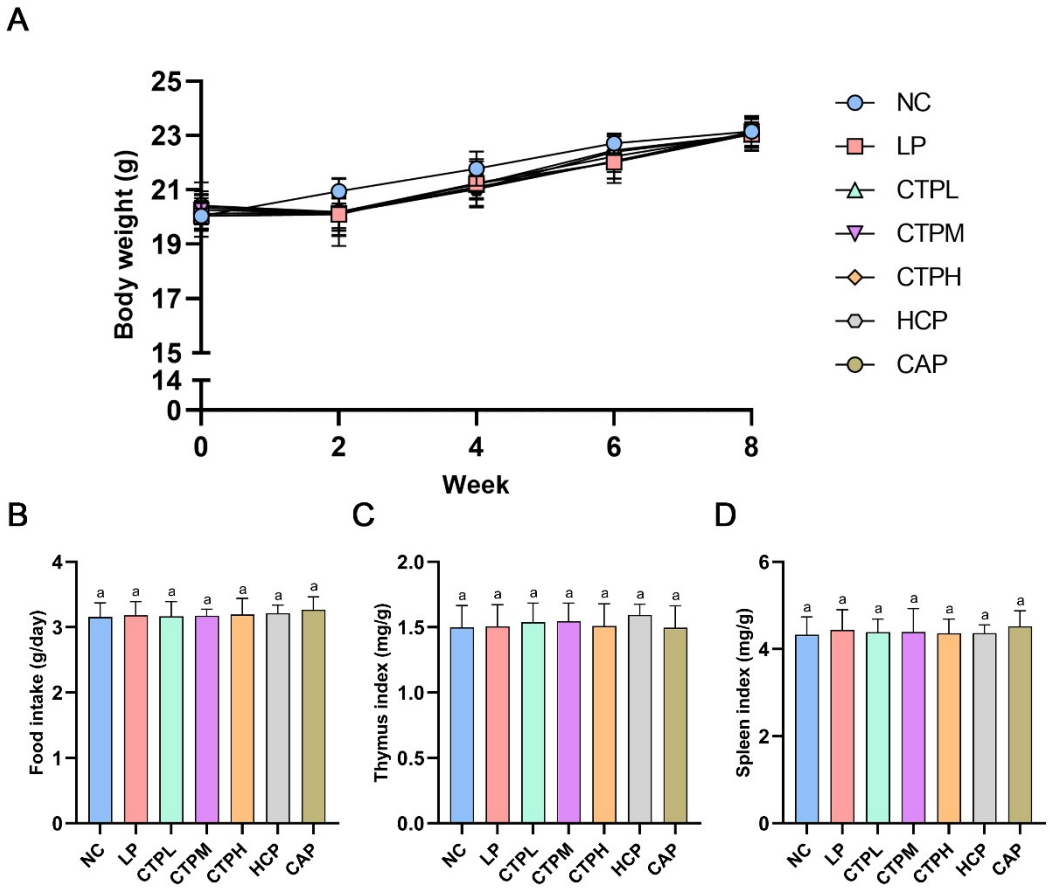


**Fig. S1** Effect of collagen peptides on body weight, food intake, and immune organ indices in mice. (A-D) Body weight, food intake, thymus index and spleen index of mice in different groups. NC, normal control group; LP, low protein group; CTP (L, M, H), hydroxyproline-containing tripeptide-rich collagen peptides low, medium and high dose groups; HCP, high molecular weight collagen peptides group; CAP, casein peptides group. Data are shown as the mean ± SEM (n = 8 per group), as determined by the one-way analysis of variance followed by Tukey's post hoc test.

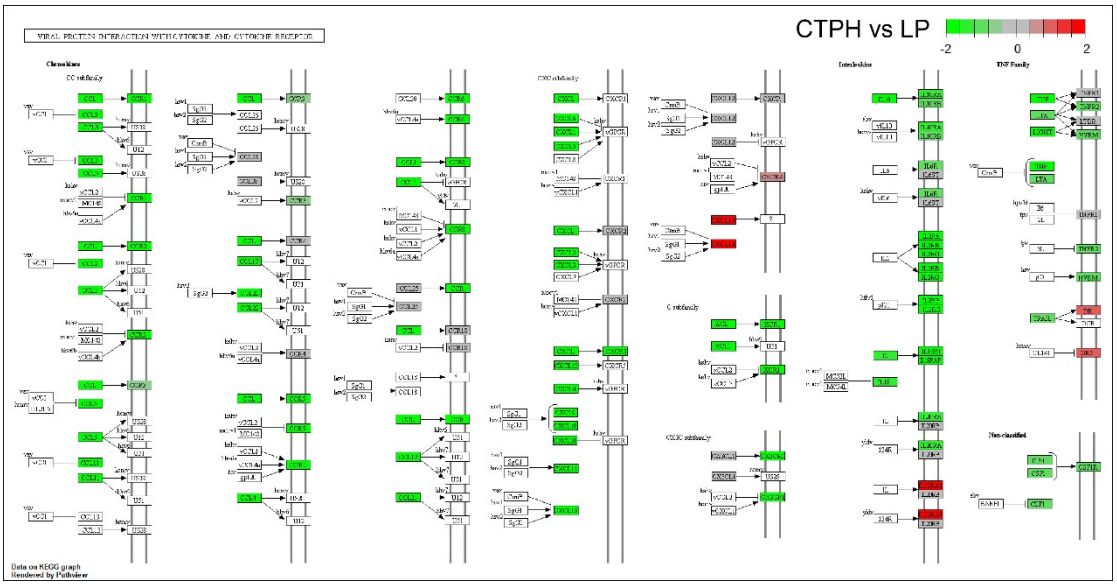


**Fig. S2** Effect of collagen peptides on cytokine signaling pathways based on KEGG enrichment analysis. KEGG pathway diagram showing viral protein interaction with cytokine and cytokine receptor. Gene information: https://www.kegg.jp/pathway/mmu04061. CTPH, hydroxyproline-containing tripeptide-rich collagen peptides high dose group; LP, low protein group.

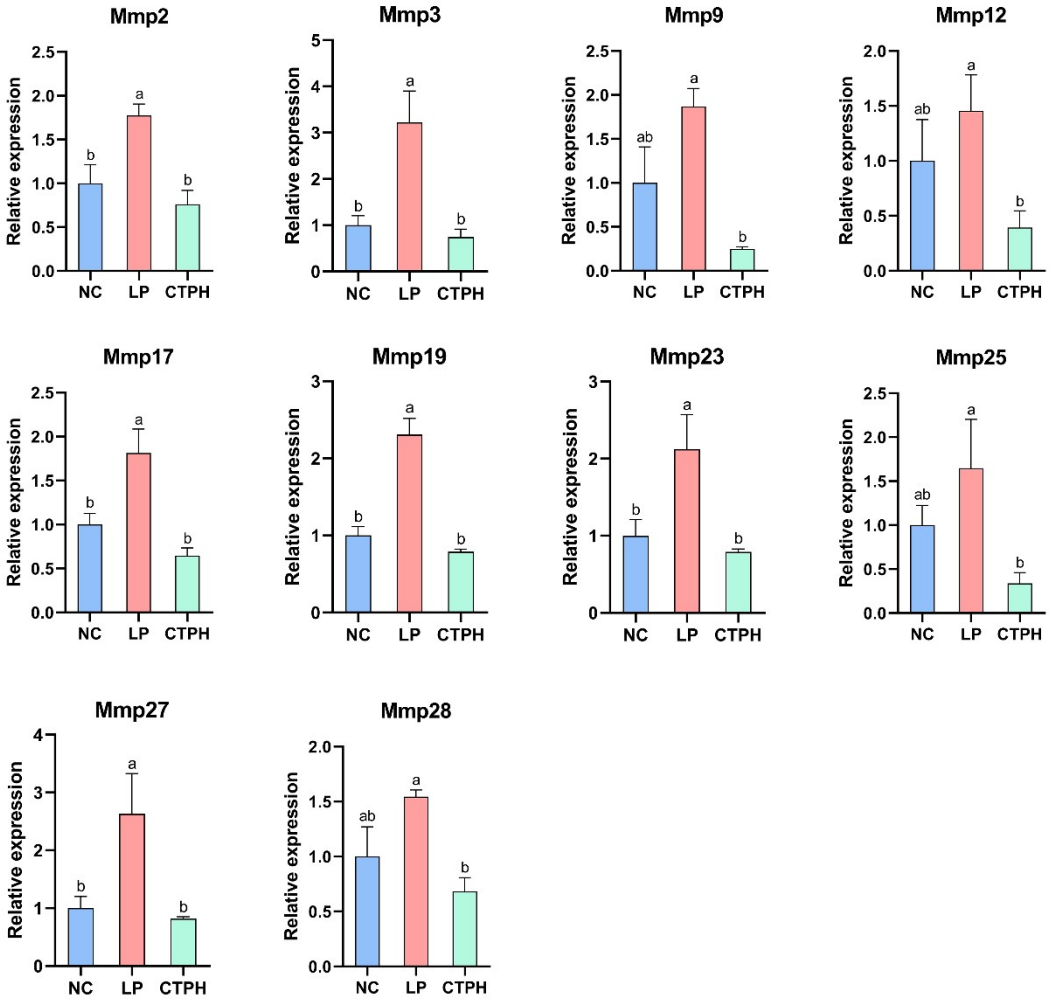


**Fig. S3** Effect of collagen peptides on the expression of matrix metalloproteinase-related genes. NC, normal control group; LP, low protein group; CTPH, hydroxyproline-containing tripeptide-rich collagen peptides high dose group. Data are shown as the mean ± SEM (n = 4 per group). Groups labeled with different letters are significantly different ($p < 0.05$), as determined by the Kruskal–Wallis test followed by pairwise Wilcoxon rank-sum tests.

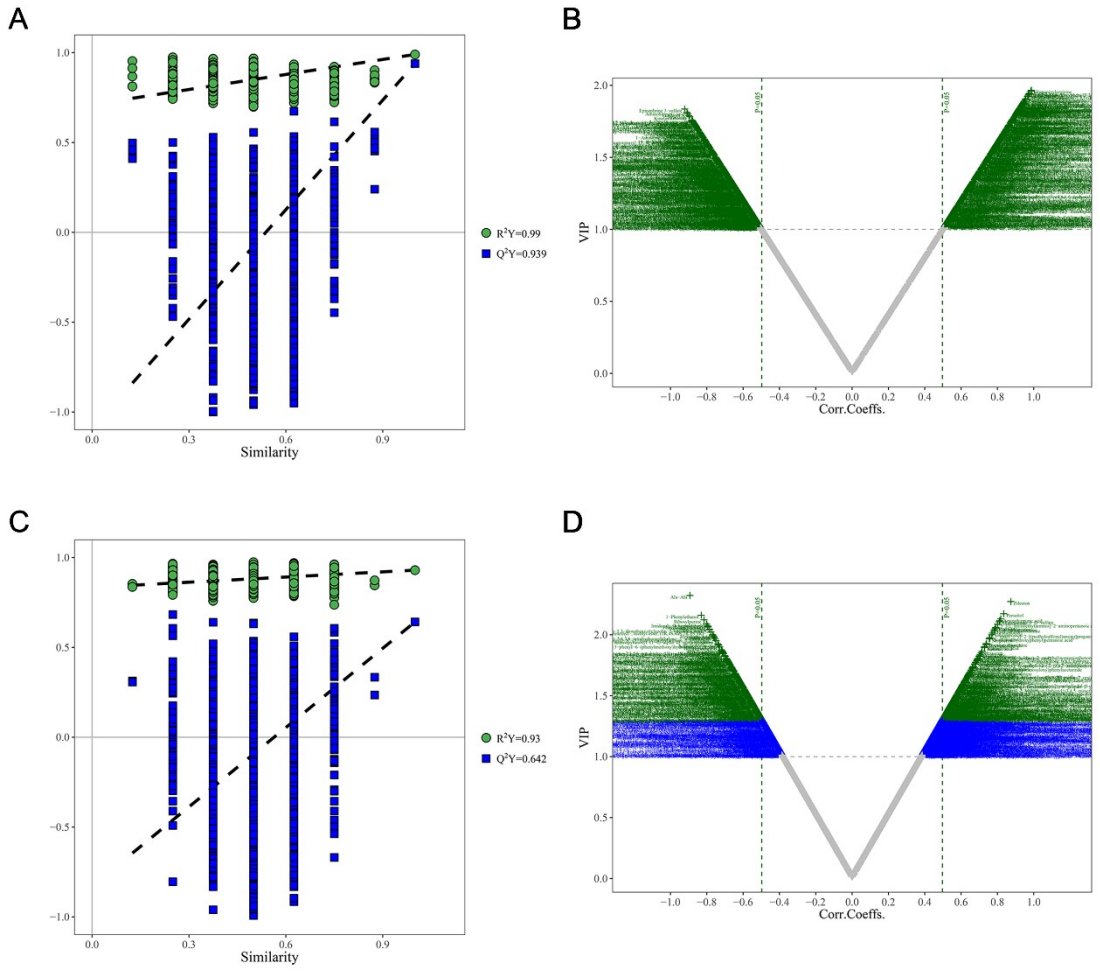


**Fig. S4** Effect of collagen peptides on fecal metabolite profiles based on OPLS-DA analysis. (A, B) Orthogonal partial least squares discriminant analysis (OPLS-DA) permutation plot and V-plot for fecal metabolite profiles comparing NC and LP groups. (C, D) OPLS-DA permutation plot and V-plot for fecal metabolite profiles comparing LP and CTPH groups (n = 6 per group). NC, normal control group; LP, low protein group; CTPH, hydroxyproline-containing tripeptide-rich collagen peptides high dose group.